

DH02: SDTM Validation Tool

Anamaria Calai

12NOV2019



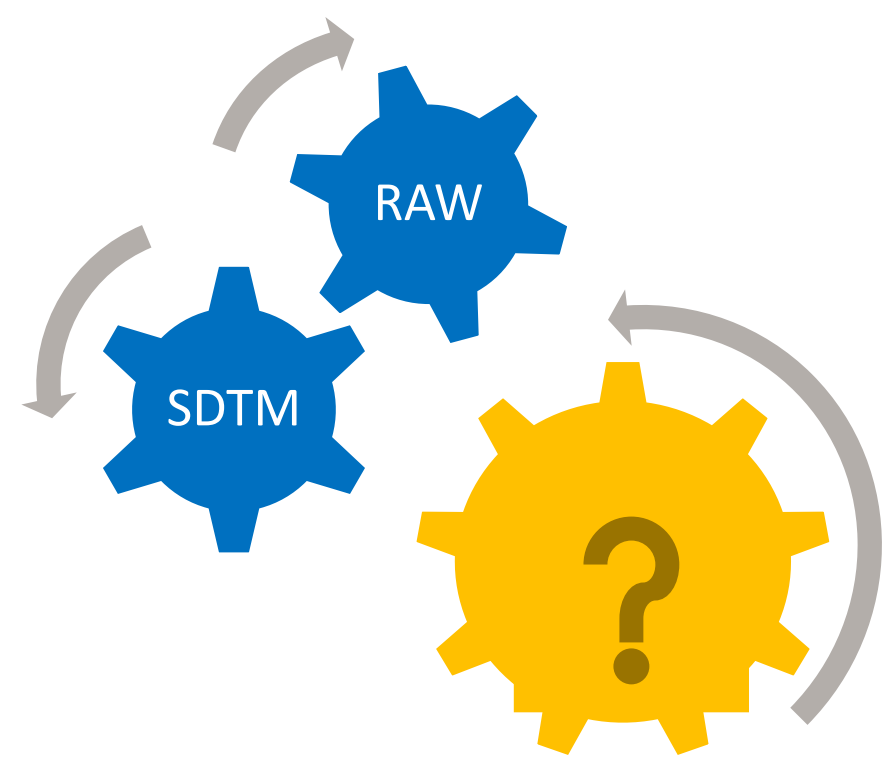
The intelligent choice
for clinical trials

SDTM Hidden Challenges

1. Do you spend a lot of time checking the consistency between RAW and SDTM data?

2. Have you struggled with data issues challenging your SV & SE programming ?

3. Do you need to manually check for sponsor specific requirements?



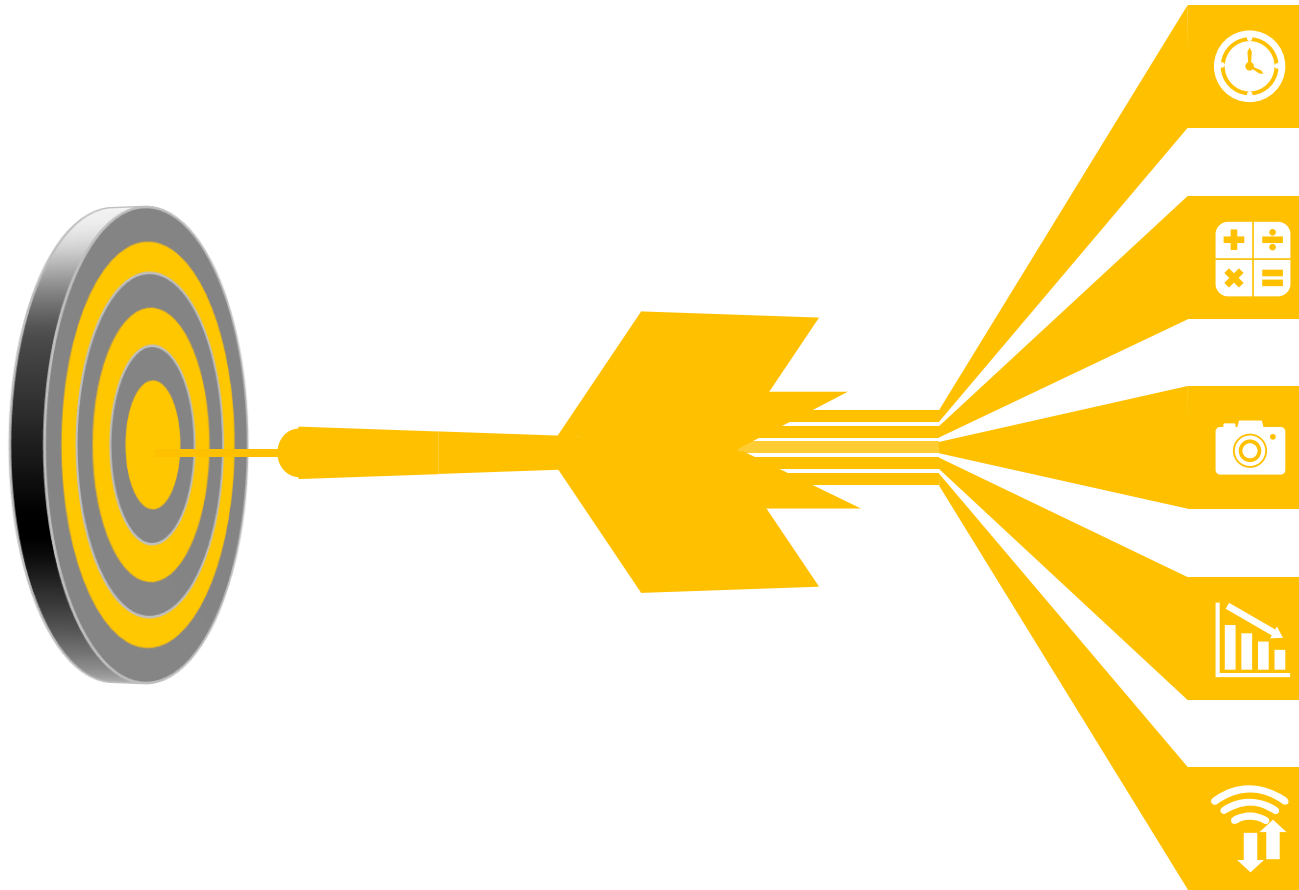
OUTPUT FROM THE VALIDATION TOOL

One single dataset containing all the identified issues consolidated



COMP FLAG	OLD_EXTRACT	NEW_EXTRACT	USUBJID	TYPE	FLAG	STATUS	COMMENT
OLD AND NEW	30AUG19:07:41:36	03SEP19:15:58:06	001	SE	Date is not in both SDTMs and SDTM.SE. Dataset=SDTM.LB Date=2019-02-18	ONGOING	Date queried with DM
OLD ISSUE, BUT FIXED	30AUG19:07:41:36		002	DV	Comparison RAW/SDTM PROC FREQ not done, as SDTM.DV is empty.	TO BE UPDATED	
NEW ISSUE		03SEP19:15:58:06	003	SV	Unscheduled date for visit 2102.01 is included already in the scheduled visit 2103		

ADDED VALUE



PROACTIVITY

Fixing issues earlier in the process of SDTM development

SPECIFICITY

Able to implement sponsor-specific checks

TRACEABILITY

Able to perform cross-checks against the RAW data

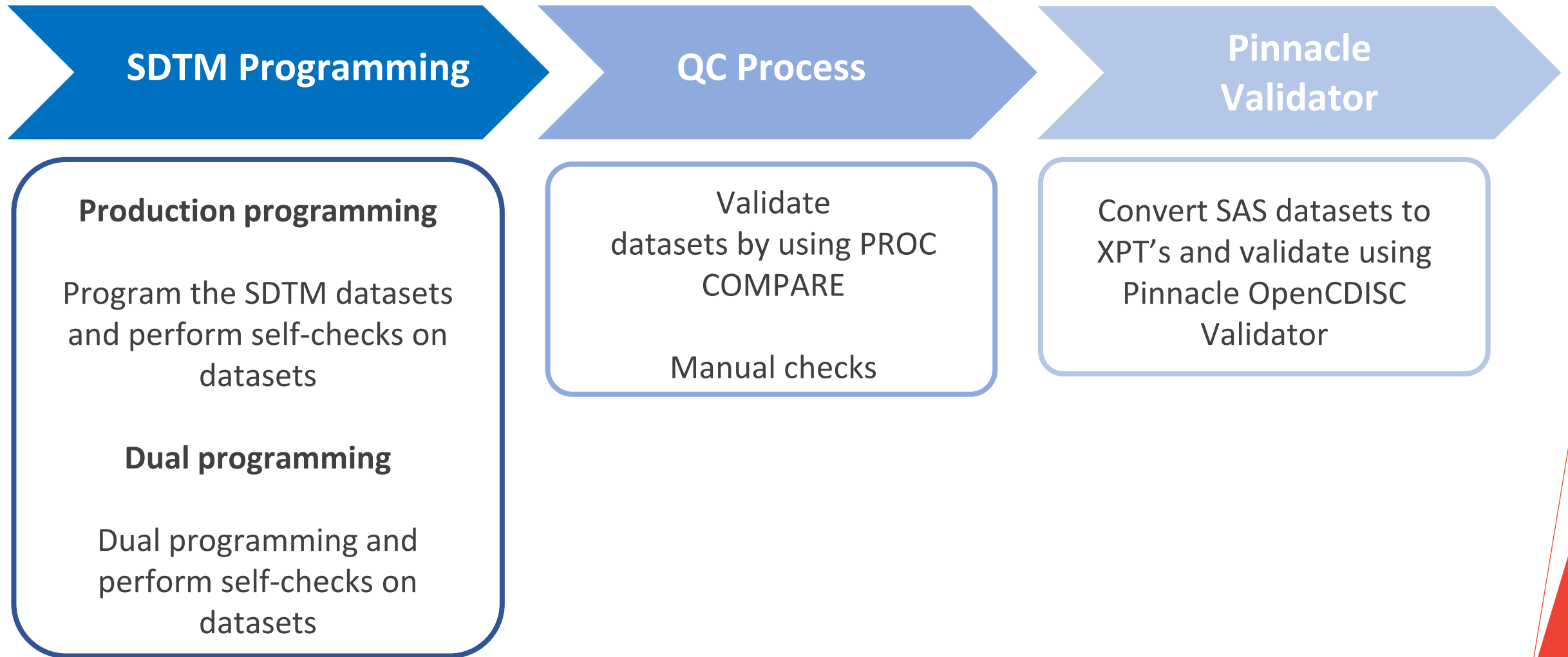
EFFICIENCY

Able to compare several reports and spot new issues more easily, but also track resolution of old issues

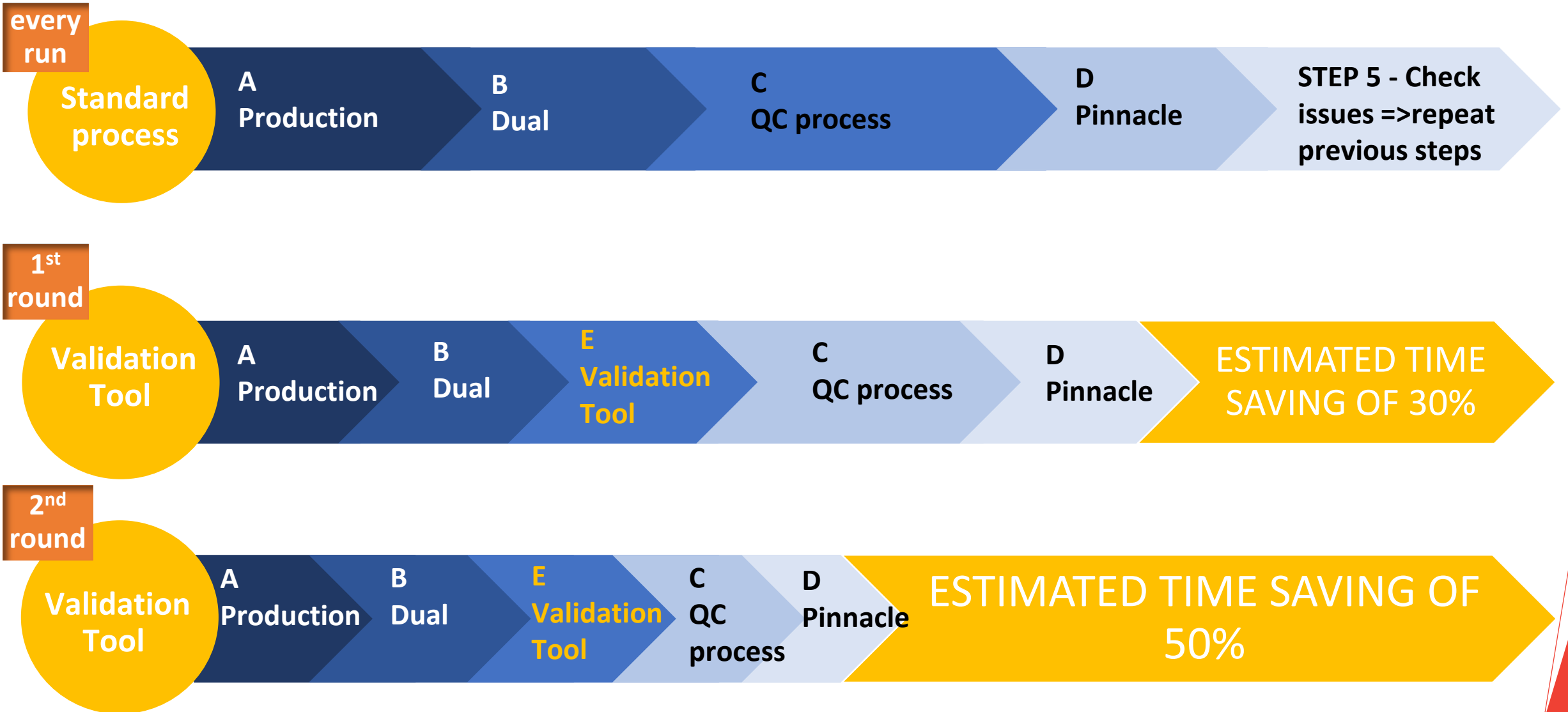
TRANSPARENCY

Having a full report of the issues from all datasets gathered in one place, accessible to any new team members

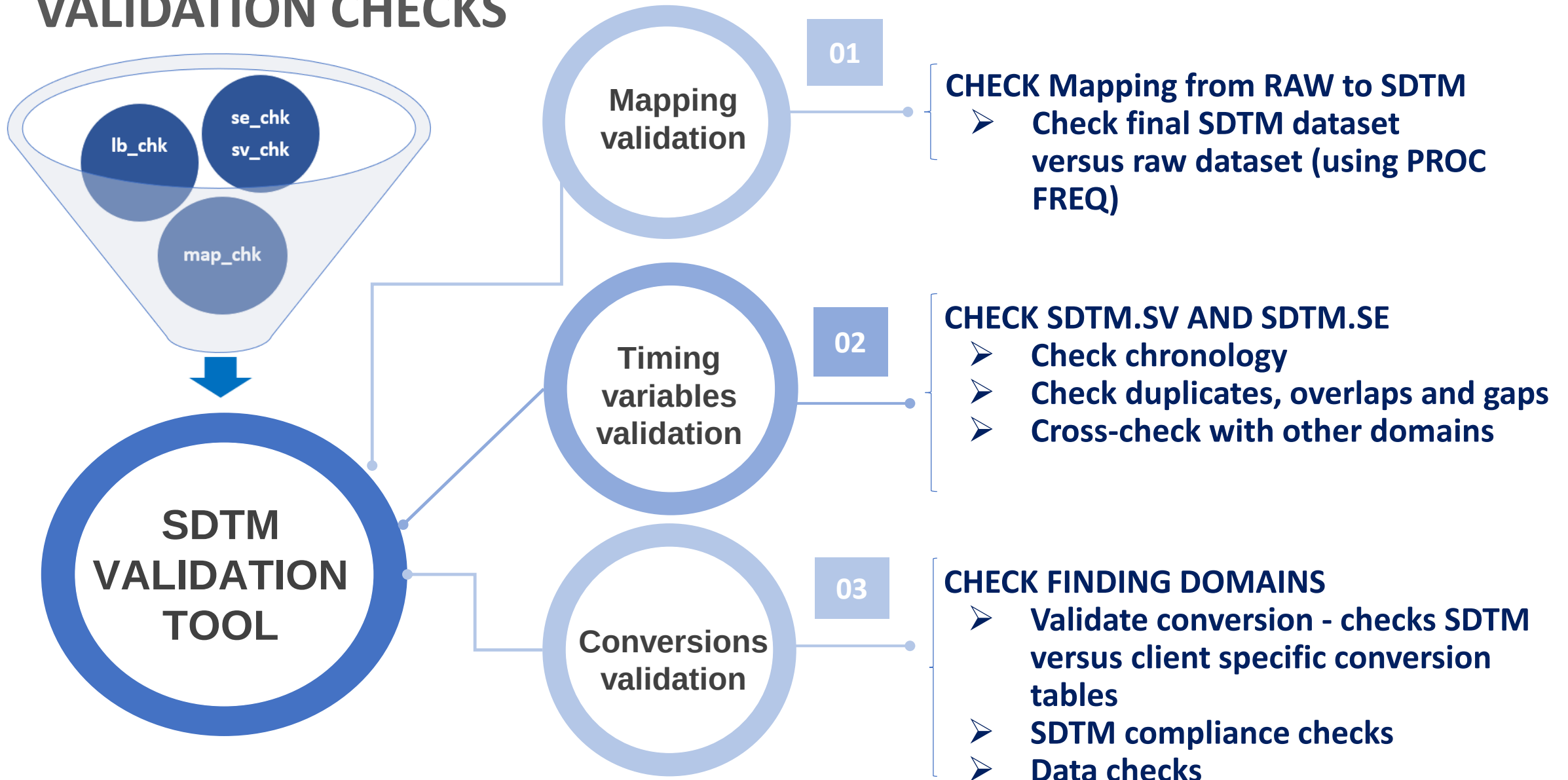
SDTM PROCESS FLOW



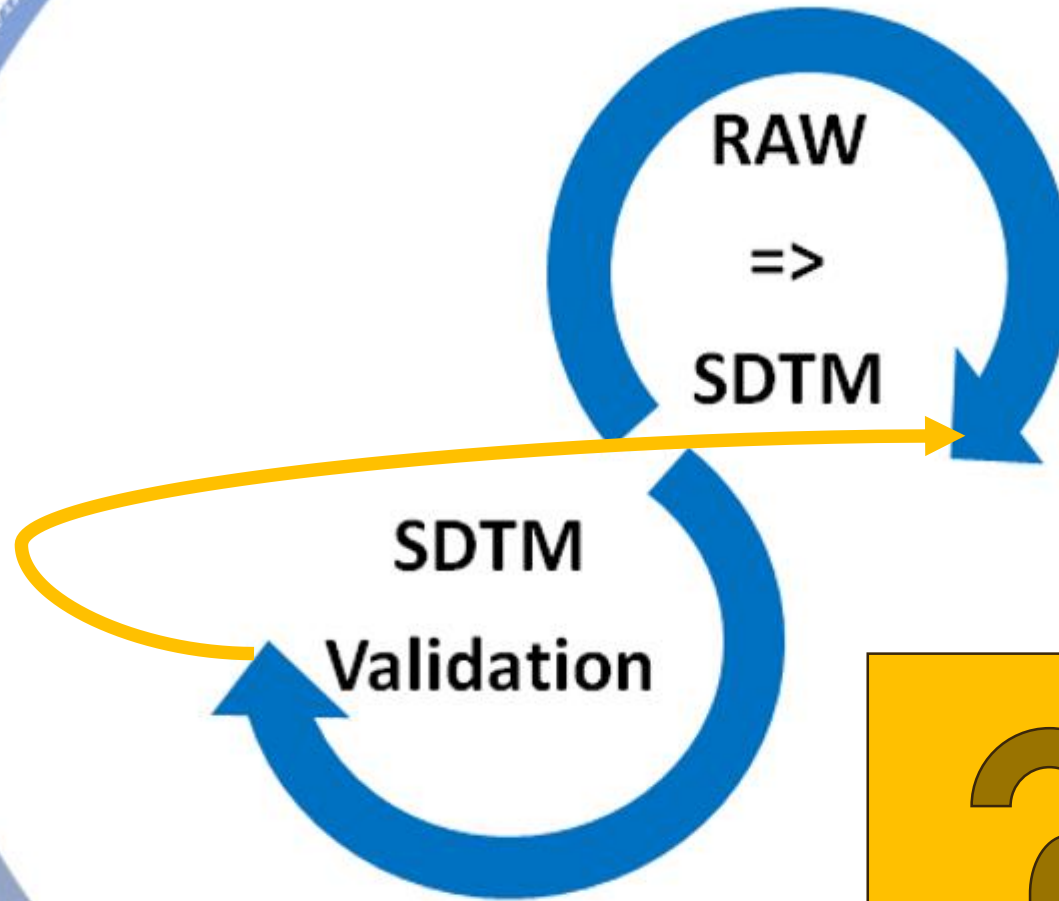
Comparison between the standard process and the new process



VALIDATION CHECKS

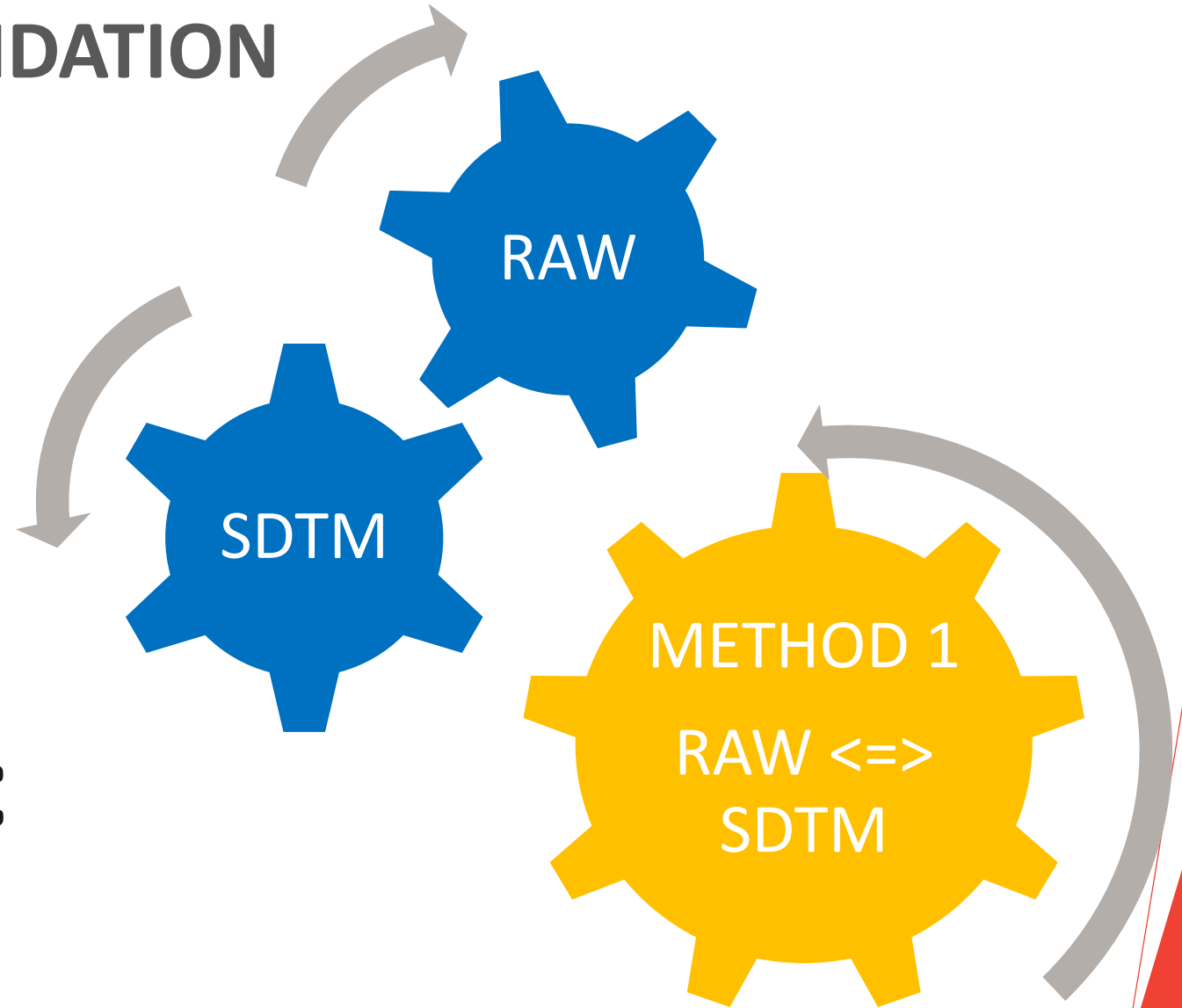


METHOD 1: MAPPING VALIDATION



METHOD 1: MAPPING VALIDATION

- What is the objective?
 - to ensure RAW/SDTM traceability
- How do I do that?
 - PROC FREQ
 - Annotated PROC CONTENTS



METHOD 1: MAPPING VALIDATION

PROC FREQ

VARIABLE	VALUE	RAW FREQUENCY
AEACN	Dose not changed	34
AEACN	Not applicable	12
AEDIS	No	46
AETERM	NEUTROPHIL COUNT DECREASED	5



VARIABLE	VALUE	SDTM FREQUENCY
AEACN	DOSE NOT CHANGED	34
AEACN	NOT APPLICABLE	12
AEDIS	N	46
AETERM	NEUTROPHIL COUNT DE	5



```
* RAW FREQUENCIES;  
ods output OneWayFreqs=rawfreq;  
proc freq data=ae_der;  
  tables _all_/ missing;  
run;  
ods output close;
```

```
* SDTM FREQUENCIES;  
ods output OneWayFreqs=sdTMfreq;  
proc freq data=work.ae_tr;  
  /* AE_TR is SDTM.AE merged with  
  SUPPAE transposed */  
  tables _all_/ missing;  
run;  
ods output close;
```

METHOD 1: MAPPING VALIDATION

VARIABLE	VALUE
AEACN	Dose not changed
AEACN	Not applicable
AEDIS	No
AETERM	NEUTROPHIL COUNT DECREASED



VARIABLE	VALUE
AEACN	DOSE NOT CHANGED
AEACN	NOT APPLICABLE
AEDIS	N
AETERM	NEUTROPHIL COUNT DE

RAW		SDTM	
RAW dataset	RAW variable	SDTM dataset	SDTM variable
AE	AEACN	AE	AEACN
AE	AE_PTXT	AE	AEDECOD
AE	AEDIS	SUPPAE	AEDIS

**Annotated
PROC
CONTENTS**

METHOD 1: MAPPING VALIDATION

VARIABLE	VALUE	FLAG	RAW FREQUENCY	SDTM FREQUENCY
AETERM	NEUTROPHIL COUNT DECREASED	Value is in RAW, but not in SDTM	5	
AETERM	NEUTROPHIL COUNT DE	Value is in SDTM, but not in RAW		5



Advantages

Links RAW and SDTM data

Unique findings

METHOD 2: TIMING DOMAINS VALIDATION

The programming of SE & SV can often be challenging, thus the need for an efficient validation is crucial.

If the standard approach is to create the SV and SE first and use these for other domains to reference, then the programming of these two domains is very important for the success of the SDTM programming.



METHOD 2: TIMING DOMAINS VALIDATION

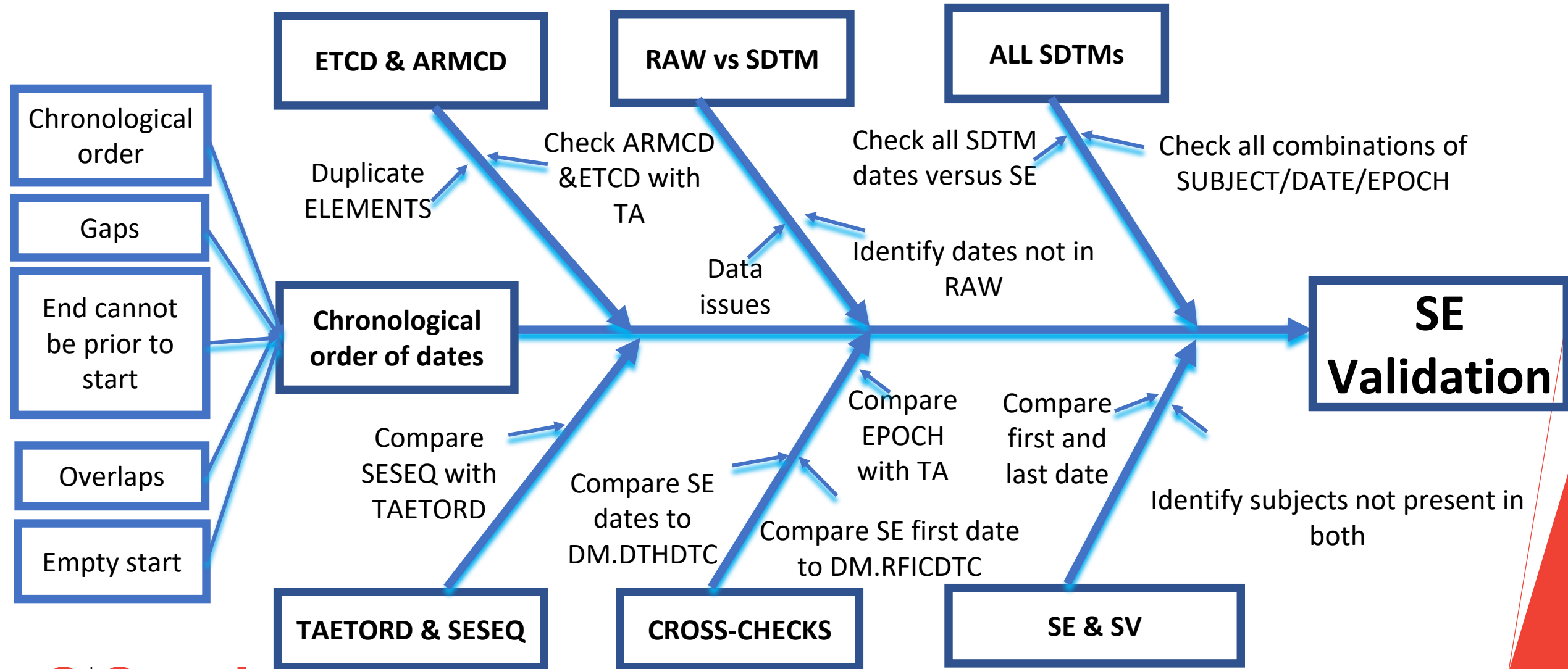
SUBJECT ELEMENTS (SDTM.SE) CHECKS

Checking the SE domain can be challenging when looking only at the final SDTM SE dataset.

Looking at the RAW data and cross-checking against the other SDTM datasets is arguably the most effective way to validate this domain.



METHOD 2: TIMING DOMAINS VALIDATION - SDTM.SE CHECKS

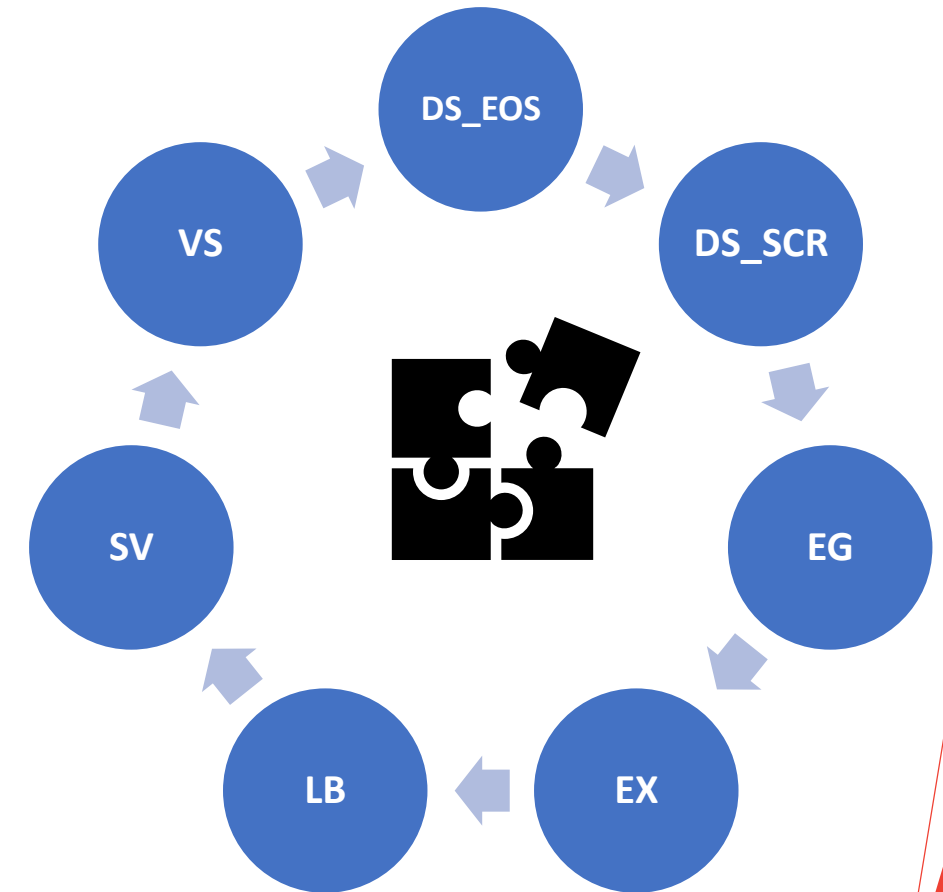


METHOD 2: TIMING DOMAINS VALIDATION - SDTM.SE CHECKS

```
data vcolumns;|
  set sashelp.vcolumn(where=(LIBNAME = "RAW" ));
run;

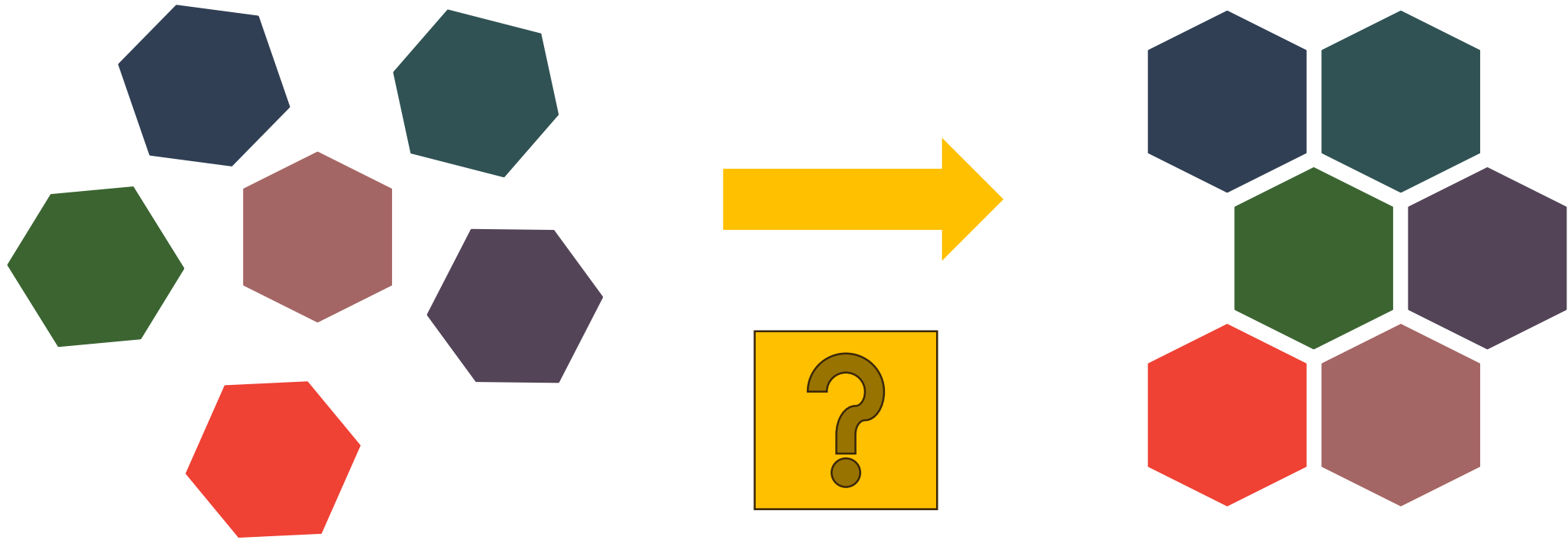
data all_dates;
set vcolumns;
  name_=compress(name, '1234567890');
  len=length(name);
  if memname not in ("AE","DD","CM","MH") and
index(name,'ENDT') eq 0 and /*check DAT*/
  ((upcase(substr(name_,length(name_)-2)) in 'DAT'));
run;

*get the list of all datasets where we want to pick
the dates from;
proc sql noprint;
  select memname into: names separated by ' ' from
all_dates;
quit;
```

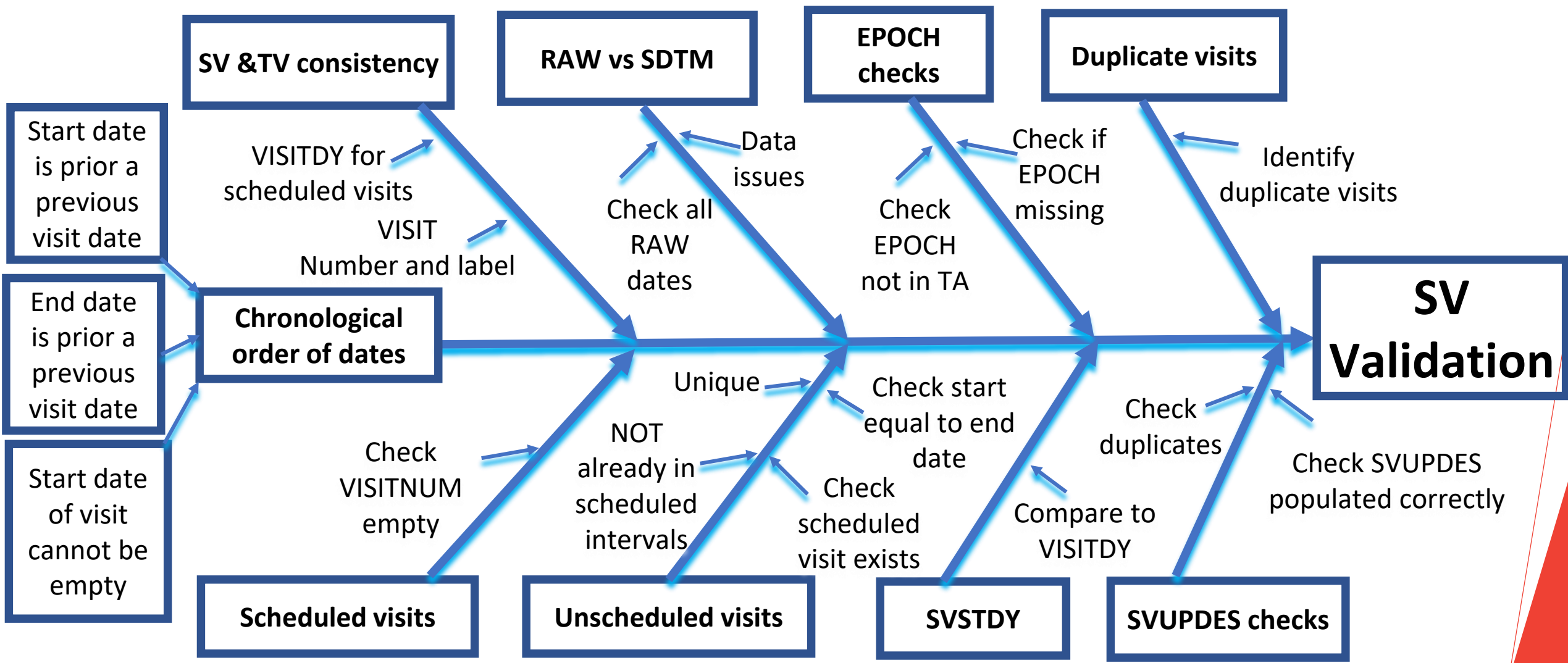


METHOD 2: TIMING DOMAINS VALIDATION - SDTM.SV CHECKS

Initial creation of an accurate SV SDTM dataset supports successful programming of all the SDTM domains, is a key reference for the overall visit timing in the study, and for identification of assessments outside the scheduled visits.



METHOD 2: TIMING DOMAINS VALIDATION - SDTM.SV CHECKS



METHOD 2: TIMING DOMAINS VALIDATION - SDTM.SV CHECKS

USUBJID	SVUPDES	VISITNUM	VISIT	VISITDY	SVSTDTC	SVENDTDC
STUDY_001		104	Day 57	57	2018-04-03	2018-05-07
STUDY_001		107.010		57	2018-04-03	2018-05-07

USUBJID	FLAG
STUDY_001	CHECK04: Unscheduled date for visit 107.01 is included already in the scheduled visit 104
STUDY_001	CHECK06: Check unscheduled visit as start and end are different 107.01 and date SVSTDTC= 2018-04-03 and SVENDTDC= 2018-05-07
STUDY_001	CHECK07: Unscheduled visit does not have SVUPDES. VISITNUM= 107.01 and date SVSTDTC= 2018-04-03 and SVENDTDC= 2018-05-07
STUDY_001	CHECK13: Planned study day VISITDY should not be populated for unscheduled visits. USUBJID=STUDY_1001051 and VISITNUM=107.01

METHOD 2: TIMING DOMAINS VALIDATION - SDTM.SV CHECKS

USUBJID	SVUPDES	VISITNUM	VISIT	VISITDY	SVSTDTC	SVENDTC
STUDY_001		104	Day 57	57	2018-04-03	2018-05-07
STUDY_001		107.010		57	2018-04-03	2018-05-07

USUBJID	FLAG
STUDY_001	CHECK04: Unscheduled date for visit 107.01 is included already in the scheduled visit 109
STUDY_001	CHECK06: Check unscheduled visit start and end are different 107.01 and date SVSTDTC= 2018-04-03 and SVENDTC= 2018-05-07
STUDY_001	CHECK07: Unscheduled visit does not have SVUPDES. VISITNUM= 107.01 and date SVSTDTC= 2018-04-03 and SVENDTC= 2018-05-07
STUDY_001	CHECK13: Planned study day VISITDY should not be populated for unscheduled visits. USUBJID=STUDY_1001051 and VISITNUM=107.01

METHOD 2: TIMING DOMAINS VALIDATION - SDTM.SV CHECKS

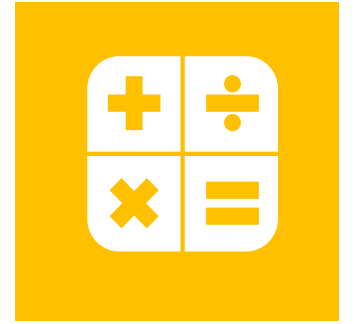
USUBJID	SVUPDES	VISITNUM	VISIT	VISITDY	SVSTDTC	SVENDTC
STUDY_001		104	Day 57	57	2018-04-03	2018-05-07
STUDY_001		107.010		57	2018-04-03	2018-05-07

USUBJID	FLAG
STUDY_001	CHECK04: Unscheduled date for visit 107.01 is included already in the scheduled visit 109
STUDY_001	CHECK06: Check unscheduled visit as start and end are different 107.01 and date SVSTDTC= 2018-04-03 and SVENDTC= 2018-05-07
STUDY_001	CHECK07: Unscheduled visit does not have SVUPDES. VISITNUM= 107.01 and date SVSTDTC= 2018-04-03 and SVENDTC= 2018-05-07
STUDY_001	CHECK13: Planned study day VISITDY should not be populated for unscheduled visits. USUBJID=STUDY_1001051 and VISITNUM=107.01

METHOD 3: CONVERSIONS Validation

The validation of the findings SDTM datasets is completed in Pinnacle based on the CDISC compliance for the domain, but this cannot account for sponsor specific rules.

Sponsor specific checks have been included in the validation tool, using the conversion tables provided by a sponsor.



METHOD 3: CONVERSIONS Validation

Sponsor specific information

TESTCD/TEST	CAT/SCAT	Measured unit	Sponsor preferred unit	Converted factor (SI)	Converted unit (SI)	PRECISION
UWBCQ / Urine white blood cells(value)	URINALYSIS		HPF			
LYM/ Lymphocytes	HEMATOLOGY / CEREBROSPINAL FLUID	Leuko/uI	cells/uL	0.001	10E9/L	0
APTT/ Activated Partial Thromboplastin Time	HEMATOLOGY / COAGULATION	sec	s	1	s	1
HOMA2IR / HOMA2-Insulin Resistance	CHEMISTRY		NO UNIT			
CODE	FLAG					
CHECK 03	SI unit is different from Reference Tables. PARM=UWBCQ TESTCD=UWBCQ LBCAT=URINALYSIS LBMETHOD= LBSPEC=URINE LBSTRESU=/uI, PREUNIT=HPF					
	LBSTRESU=10E9/L, CNVFCR=0.001, PCSNUM=2, RESC=2.72, LBSTRESC=2717.00, LBORRES=2717					
CHECK05	Precision is different from Reference Tables, please check PARM=APTT LBTESTCD=APTT PCSNUM=1 LBSTRESU=s SDTM_PREC=0					
CHECK07	Check if result is measurable, as the unit suggests we should not have numeric results. TESTCD=HOMA2IR LBCAT=CHEMISTRY LBMETHOD= LBSPEC=PLASMA/SERUM LBORRES=6.37 LBSTRESU=NO UNIT					

METHOD 3: CONVERSIONS Validation

Sponsor specific information

TESTCD/TEST	CAT/SCAT	Measured unit	Sponsor preferred unit	Converted factor (SI)	Converted unit (SI)	PRECISION
UWBCQ / Urine white blood	URINALYSIS		HPF			
LYM/ Lymphocytes	HEMATOLOGY / CEREBROSPINAL FLUID	Leuko/ul	cells/uL	0.001	10E9/L	0
APTT/ Activated Partial Thromboplastin Time	HEMATOLOGY / COAGULATION	sec	s	1	s	1
HOMA2IR / HOMA2-Insulin Resistance	CHEMISTRY		NO UNIT			
	CODE	FLAG				
	CHECK03	SI unit is different from Reference Tables. PARM=UWBCQ TESTCD=UWBCQ LBCAT=URINALYSIS LBMETHOD= LBSPEC=URINE LBSTRESU=/ul ,				
	CHECK04	Conversion is not correct (compare re-derived RESC with SDTM.LBSTRESC), please check PARM=LYM, LBTESTCD=LYM, PREF_UNIT=Leuko/ul, LBSTRESU=10E9/L, CNVFCR=0.001, PCSNUM=2, RESC=2.72, LBSTRESC=2717.00 , LBORRES=2717				
	CHECK05	Precision is different from Reference Tables, please check PARM=APTT LBTESTCD=APTT PCSNUM=1 LBSTRESU=SDTM_PRESU				
	CHECK07	Check if result is measurable, as the unit suggests we should not have numeric results. TESTCD=HOMA2IR LBCAT=CHEMISTRY LBMETHOD= LBSPEC=PLASMA/SERUM LBORRES=6.37 LBSTRESU=NO UNIT				

METHOD 3: CONVERSIONS Validation

Sponsor specific information

TESTCD/TEST	CAT/SCAT	Measured unit	Sponsor preferred unit	Converted factor (SI)	Converted unit (SI)	PRECISION
UWBCQ / Urine white blood cells(value)	URINALYSIS		HPF			
LYM/ Lymphocytes	HEMATOLOGY / CEREbroSPINAL FLUID	Leuko/ul	cells/uL	0.001	10E9/L	0
APTT/ Activated Partial Thromboplastin Time	HEMATOLOGY / COAGULATION	sec	s	1	s	1
Resistance						
CODE	FLAG					
CHECK03	SI unit is different from Reference Tables. PARM=UWBCQ TESTCD=UWBCQ LBCAT=URINALYSIS LBMETHOD= LBSPEC=URINE LBSTRESU=/ul , PREUNIT=HPF					
CHECK04	Conversion is not correct (compare re-derived RESC with SDTM.LBSTRESC), please check PARM=LYM, LBTESTCD=LYM, PREF_UNIT=Leuko/ul, LBSTRESU=10E9/L, CNVFCR=0.001, PCSNUM=2, RESC=2.72, LBSTRESC=2717.00 , LBORRES=2717					
CHECK 05	Precision is different from Reference Tables, please check PARM=APTT LBTESTCD=APTT PCSNUM=1 LBSTRESU=s SDTM_PREC=0					
	LBSPEC=PLASMA/SERUM LBORRES=6.37 LBSTRESU=NO UNIT					

METHOD 3: CONVERSIONS Validation

Sponsor specific information

TESTCD/TEST	CAT/SCAT	Measured unit	Sponsor preferred unit	Converted factor (SI)	Converted unit (SI)	PRECISION
UWBCQ / Urine white blood cells(value)	URINALYSIS		HPF			
LYM/ Lymphocytes	HEMATOLOGY / CEREBROSPINAL FLUID	Leuko/ul	cells/uL	0.001	10E9/L	0
APTT/ Activated Partial	HEMATOLOGY / COAGULATION	sec	s	1	s	1
HOMA2IR / HOMA2-Insulin Resistance	CHEMISTRY		NO UNIT			

CODE	FLAG
CHECK03	SI unit is different from Reference Tables. PARM=UWBCQ TESTCD=UWBCQ LBCAT=URINALYSIS LBMETHOD= LBSPEC=URINE LBSTRESU=/ul , PREUNIT=HPF
CHECK04	Conversion is not correct (compare re-derived RESC with SDTM.LBSTRESC), please check PARM=LYM, LBTESTCD=LYM, PREF_UNIT=Leuko/ul, LBSTRESU=10E9/L, CNVFCR=0.001, PCSNUM=2, RESC=2.72, LBSTRESC=2717.00 , LBORRES=2717
CHECK05	Precision is different from Reference Tables. please check PARM=APTT LBTESTCD=APTT PCSNUM=1 LBSTRESU=s SDTM PREC=0
CHECK 07	Check if result is measurable, as the unit suggests we should not have numeric results. TESTCD=HOMA2IR LBCAT=CHEMISTRY LBMETHOD= LBSPEC=PLASMA/SERUM LBORRES=6.37 LBSTRESU=NO UNIT

Any Questions?

Anamaria Calai

acalai@cmedresearch.com